

University of Groningen

Performance of the ICU

Moreno, Rui Paolo Jinó

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

1997

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Moreno, R. P. J. (1997). *Performance of the ICU: are we able to measure it?* s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Chapter 4

Length of stay in the ICU, case mix and severity of illness

Rui Moreno¹, Dinis Reis Miranda²

¹ Intensive Care Unit, Hospital de Santo António dos Capuchos, Lisboa, Portugal

² ICU Research Group, Department of Surgery, University Hospital of Groningen,
the Netherlands

Submitted for publication

INTRODUCTION

In the last 15 years a large effort has been made to build and improve general outcome prediction models. The models were developed on large databases of intensive care patients. The important characteristic of the models is that they used data collected at admission to the ICU [1] or during the first 24 hours after admission in the ICU [2-4]. Outcome prediction is computed after using advanced statistical techniques to select and attribute weights to the variables in the collected data. These models have been used in the evaluation of the performance of intensive care [5-8], in the comparison of the ICUs [9-11] and to assess comparability of groups in clinical trials [12-15].

More recently, the applicability of the models to populations, which differs from the ones where they have been developed and validated, was challenged [16,17]. The exact extension to which different patient mix characteristics (*e.g.* type of patient, location in the hospital before ICU admission, diagnostic category of admission) affect the performance of general outcome prediction models is not completely known but it is assumed to be significant [16-19].

Another potential problem is that the models used today focus on the prediction of the outcome at hospital discharge. The rationale behind this choice is that hospital mortality can be well defined, is a very relevant outcome and it minimises the impact of different policies of discharge at ICU level when comparing outcomes. However, the choice of a precise point in time rather than the evaluation of survival curves, can introduce analytic difficulties in those cases when the binary outcome of interest (survivor/non-survivor) is very far away in time from the moment when data was collected for prediction. It has been shown that the determinants of mortality change over time and that models based on clinical data collected at admission or at 24 hours in the ICU, can loose their prediction accuracy with time [20,21]. This rational led several authors [20-25] to propose the use of data collected at different points in time to adjust predictions made shortly after admission.

The extent to which the performance of a model developed and calibrated upon first 24 hours data is affected by the length of stay (LOS) in the ICU has not been formally tested. It is generally assumed that the model applies to all treated patients, independently of their LOS in the ICU. This assumption is usually applicable since about half the patients leave the ICU within 48 hours of admission and only 14 % stay more than three weeks in the ICU [26]. This explains perhaps why the most (though not all) populations where the models were found not to accurately predict mortality, were generally associated with particular conditions with longer LOS [27-30]. If the ICUs deal with various aggregates of patient baseline characteristics known to have a significant impact on LOS (medical patients or patients admitted from the ward, other ICU or other hospital), or if the mortality rate is high, the different performance of the models across different LOS groups can introduce bias in the comparison of predicted and observed mortality among the ICUs. As a matter of fact, the

differences of clinical performance at ICU level might be explained only by the application of a general model to a sub-set of patients in which the model does not perform with sufficient accuracy.

The aim of the present study is threefold:

- to analyse the relations between patient baseline characteristics and LOS;
- to explore the influence of two hospital characteristics (type of hospital and the presence of step-down units) on LOS;
- to analyse the effects of LOS on the predictive performance of the new Simplified Acute Physiology Score (SAPS II) [4].

MATERIALS AND METHODS

THE DATABASE

The study used the database of the Foundation for Research on Intensive Care in Europe (FRICE) concerning data collected during a concerted action of the Biomed 1 program of the Commission of the European Communities, called EURICUS-I (BMH1-CT93-1340), addressing the effects of organisation and management on the effectiveness and efficiency of ICUs in Europe, and which methods and results will be published elsewhere, including the collection of medical data and non-medical data on 89 ICUs of 12 countries of Europe.

In each country, the selection of the ICUs was made to include units of hospitals of different types (university and non-university) and different sizes. In other words, in order to generate the largest possible variation of organisation and management variables, the ICUs were selected so that at least each country would participate with 4-8 ICUs operating at different levels of care [31].

PATIENT DATA

Data were collected during a 4-month period (October 3, 1994 to January 31, 1995) on 16060 patients, consecutively admitted to the ICUs in EURICUS-I. The exclusion criteria for SAPS II followed those expressed in the original description [4]: patients younger than 18 years of age, readmission's, acute coronary care patients, burns and patients in the post-operative period after coronary artery by-pass surgery. The exclusion of the patients with less than 8 hours in the ICU was not done in order to evaluate the variations in performance of SAPS II according to the length of stay.


For each patient a simple set of variables was collected including all the variables described

for the original SAPS II system. Basic demographic characteristics including age, location in the hospital before admission to the ICU, type of patient (medical and scheduled/unscheduled surgical), were also collected as well as the principal diagnostic category on admission at the ICU using a list of 78 mutually exclusive diagnoses [3]. Basic characteristics of the ICU and of the hospital were also evaluated.

Patients were followed up to the hospital discharge and their survival status was then registered. The patients still in the hospital or with unknown outcome on May 1, 1995 (three months after the end of data collection) were dropped from the study (113 patients).

Informed consent of the patients or next of kin was considered not necessary and was waived by the Institutional Review Board of the participating hospitals.

METHODS

The evaluation of the overall goodness of fit of SAPS II was done by formal testing of discrimination and calibration. Discrimination was evaluated by the total area under the ROC curve, computed by a modification of the Wilcoxon statistics, as proposed by Hanley and McNeil [32]. Calibration was assessed by Hosmer-Lemeshow goodness-of-fit test , collapsing the table based on deciles of the estimated probabilities, and comparing the observed versus expected number of patients in each of the 20 cells to determine whether the discrepancies are acceptably small [33,34]. Observed/expected (O/E) mortality ratios were computed dividing the observed number of deaths by the number of deaths estimated by the model. Confidence intervals for this ratio were computed using a parametric approach, as described by Rapoport et al. [35].

In the evaluation of the relations between the LOS and severity of illness stratified analyses were performed after dividing the patients according to their LOS in the ICU. The performance of SAPS II was then assessed in each of these mutually exclusive subgroups using the techniques described above.

In a further step a regression model was developed, relating severity of illness to LOS. We considered in this analysis all the patients with a SAPS II between the 2.5 and the 97.5 percentile. In order to increase the number of patients in each SAPS II group, we aggregate SAPS II score at 5 point intervals. Then, the database was spliced randomly into development (2/3 of the patients) and validation samples (1/3 of the patients) and a model relating severity of illness to mean LOS in each group was estimated in the development sample and its results cross-validated into the validation sample.

The evaluation of the impact of patient related characteristics on LOS was done stratifying patients according to severity of illness (SAPS II), type of patient (medical, scheduled surgery and unscheduled surgery), location in the hospital before ICU admission (operative theatre/recovery room, emergency room, ward and other ICU/other hospital) and diagnostic category of admission (using a list of 78 mutually exclusive diagnoses [3]).

The evaluation of the impact of hospital related characteristics on LOS was done stratifying

patients according to type of hospital (< 300 beds, 300-500 beds, > 500 beds and university) and to the presence or absence in the same hospital of step-down units. In 9 hospitals (10 %) the presence or absence of step-down units could not be assessed and they were excluded from the respective analysis.

Chi-square statistics were used to test for the statistical significance of categorical variables and one-way analysis of variance was used to assess continuous variables. All statistical tests were two-sided, and a significance level of 0.05 was used except when stated otherwise. In the comparison of the LOS we used Mann-Whitney U test (two groups comparison) or Kruskal-Wallis H test (multiple groups comparison), since the distribution was highly skewed. Results are presented as median (interquartile range) except when stated otherwise.

Data analysis and statistics were performed using the Statistical Package for Social Sciences (SPSS) version 6.0.1 for Windows at the University Hospital of Groningen, the Netherlands.

RESULTS

The 89 participating ICUs comprised a diverse sample of ICUs, in 12 countries of Europe, with 35 (39 %) being in university hospitals. Median hospital size was 534 beds, with 39 % of the ICU beds in hospitals with less than 500 beds, 50 % in hospitals with 500 - 1000 beds, and 11 % in hospitals with more than 1000 beds. Median number of beds per ICU was 8 (interquartile range 6 to 10 beds). During the study period, data were collected on 16060 consecutive admissions.

After applying the patient exclusion criteria described above, the data of 11574 patients were available to the study (72.1 %): 9114 survivors (78.7 %) and 2460 non-survivors (21.3 %). The median number of analysed patients per ICU was 109 (interquartile range 75-161). Basic characteristics of the studied population are described in Table 1. Median LOS was 2.1 days (interquartile range 0.9 to 5.7 days), significantly longer ($p < 0.001$) in non-survivors (median 3.4 days, interquartile range 1.0 to 10.4 days) than in survivors (median 1.9 days, interquartile range 0.9 to 4.8 days) (Figure 1). Median LOS presented very large variations between ICUs, with values ranging from 0.9 days (interquartile range 0.8 to 1.4 days) to 11.8 days (interquartile range 4.2 to 26.6 days).

Table 1. Basic descriptive characteristics of the studied population.

Countries	12
European areas	13
Number of ICUs	89
Number of analysed patients	11574
Number of patients per ICU ^a	109 (75-161)
Age, years (mean \pm SD)	59.4 \pm 17.2
Type of Patient (%)	
Medical	56.8
Surgical, scheduled	23.3
Surgical, unscheduled ^b	19.9
Mechanical ventilation (%) ^a	41.4
SAPS II score	31 (21-44)
SAPS II probability of hospital mortality ^a	11.7 (4.2-32.6)
ICU mortality (%)	15.0
LOS in the ICU, days ^a	2.1 (0.9-5.7)
Hospital mortality (%)	21.3

ICU, Intensive care Unit; SD, standard deviation; LOS, length of stay; SAPS II, new Simplified Acute Physiology Score.

^a: median, (interquartile range)

^b: during the first 24 hours in the ICU

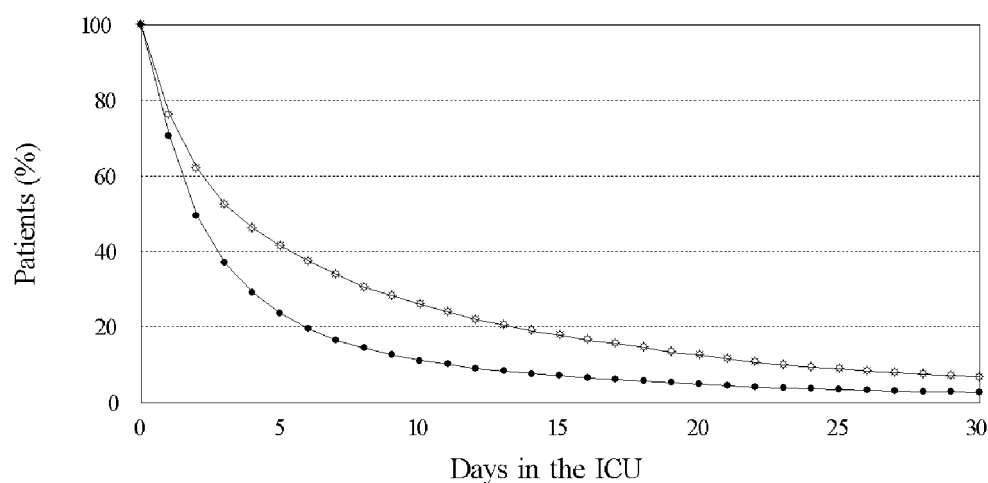


Figure 1. Hospital outcome and length of stay in the ICU. Data are presented as percentage of patients still in the intensive care unit (ICU). Non-survivors (n = 2460) (*) presented a significantly longer ($p < 0.001$) length of stay in the ICU than survivors (n = 9114) (•).

LOS AND TYPE OF PATIENT

The LOS was longer in medical patients (median 2,9 days, interquartile range 1.2 to 6.7 days) than in unscheduled surgery patients (median 2.5 days, interquartile range 0.9 to 7.0 days) and scheduled surgery patients (median 1.0 days, interquartile range 0.9 to 2.6 days). The difference of LOS between medical/unscheduled surgery patients and scheduled surgery patients was significant ($p < 0.001$).

LOS AND LOCATION IN THE HOSPITAL BEFORE ADMISSION TO THE ICU

Location in the hospital prior to ICU admission presented a significant impact on LOS ($p < 0.001$). LOS was longer for patients admitted from other ICUs/other hospitals (median 4.0 days, interquartile range 1.7 to 10.9 days), intermediate for patients admitted from the ward (median 3.2 days, interquartile range 1.3 to 7.7 days) and from the emergency room (median 2.3 days, interquartile range 1.0 to 5.0 days), and shorter for patients admitted from the operative theatre/recovery room (median 1.2 days, interquartile range 0.9 to 3.1 days).

LOS AND DIAGNOSTIC CATEGORY OF ADMISSION

Within each diagnostic category of admission, the LOS in the ICU presented a significant relation with the surgical status of the patient in cardiovascular, respiratory, neurological and other diagnostic categories ($p < 0.001$); for gastrointestinal and trauma diagnostic categories no such relation could be demonstrated (Table 2). Inside the same operative status, differences were also apparent between different diagnostic categories, with non-operative respiratory and trauma (non-operative and post-operative) patients having the longer LOS.

Table 2. Diagnostic category of admission (3) and length of stay in the ICU.

Diagnostic category of Admission	Non-operative		Post-operative		<i>p</i>
	No.	Median (IQ)	No.	Median (IQ)	
Cardiovascular	2173	2.6 (1.2-4.9)	805	1.4 (0.9-3.0)	< 0.001
Respiratory	1608	4.5 (1.8-10.4)	542	1.2 (0.9-3.1)	< 0.001
Gastrointestinal	423	2.1 (1.0-5.5)	1771	2.0 (0.9-5.3)	NS
Neurological	720	2.9 (1.0-7.8)	617	1.0 (0.8-2.9)	< 0.001
Trauma	347	4.3 (1.8-13.5)	432	4.1 (1.4-11.7)	NS
Other	1456	1.8 (0.8-4.3)	680	0.9 (0.8-1.7)	< 0.001

No., number of patients; IQ, interquartile range, NS, non significant.

LOS AND SEVERITY OF ILLNESS

Median SAPS II score was 31 (interquartile range 21 to 44) and median SAPS II probability of hospital mortality was 11.7 % (interquartile range 4.2 to 32.6 %). In the analysed sample the area under the ROC curve for SAPS II was 0.816 (standard error 0.005). Hosmer-

Lemeshow goodness-of-fit test C was 284.5 ($p < 0.001$) for SAPS II, implying that the combined discrepancy between observed and expected outcomes in survivors and in non-survivors was out of the sampling variance or, in other words, that the model did not accurately predict the outcome of the patients in this cohort. Overall O/E mortality ratio was 0.95 (95 % confidence interval 0.92 to 0.98).

SAPS II predicted risk of death, computed in the first 24 hours of admission was significantly higher in patients with longer stays in the ICU, with values ranging from 16.6 % for patients with a LOS less than 1 day to 45.8 % for patients with a LOS greater than one month. However, the relation between SAPS II score and LOS was not linear, with lower LOS in both extremes of severity (very low and very high SAPS II), as shown in Figure 2. This non linearity could be explained by different relationship between SAPS II score and LOS in survivors and non-survivors (Figure 2). SAPS II was able, in this sample to explain 66 % in the variations of the LOS.

The performance of SAPS II in this population was substantially affected by the length of stay in the ICU of the analysed patients, both in its discriminative power and on calibration. As presented in Figure 3, the area under ROC curve presented a significant relation with the LOS, with values ranging from 0.888 (standard error 0.009) for patients remaining in the ICU less than 1 day to 0.594 (standard error 0.018) for patients with a LOS greater than two weeks. Calibration was also affected by the LOS. As presented in Table 3 Hosmer-Lemeshow test C and O/E mortality ratios were both affected by LOS, mainly in the patients with longer stays. The effect of these findings will depend on the particular composition of the population admitted to the ICU, since the percentage of patients with a LOS greater than 7 days varied at ICU level from 5.2 % to 63.2 % of the admitted patients and the percentage of patients with a LOS greater than 15 days varied at ICU level from 0 % to 42.9 % of the admitted patients.

Table 3. Hosmer-Lemeshow goodness-of-fit test C and observed/expected mortality ratios according to length of stay in the ICU (LOS).

LOS (days)	No.	Chi-square	<i>p</i>	Observed Deaths	Predicted Deaths	Mortality ratio (95 % confidence intervals)
0-	3267	50.03	< 0.001	585	543	1.08 (1.02-1.13)
1-	2278	36.14	< 0.001	348	403	0.86 (0.79-0.93)
2-	1366	22.32	0.004	235	282	0.83 (0.75-0.92)
3-	882	33.12	< 0.001	158	192	0.82 (0.72-0.93)
4-	621	46.02	< 0.001	113	146	0.78 (0.66-0.89)
5-	459	47.87	< 0.001	102	122	0.84 (0.71-0.96)
6-	365	34.19	< 0.001	85	102	0.83 (0.69-0.97)
7-	276	39.16	< 0.001	86	87	0.99 (0.84-1.13)
8-14	981	106.74	< 0.001	311	317	0.98 (0.91-1.06)
≥ 15	1079	457.94	< 0.001	437	391	1.12 (1.05-1.18)

No., number of patients

Figure 2. Length of stay in the ICU according to the new Simplified Acute Physiology Score

(SAPS II) score. Top: survivors; middle, non-survivors; bottom: overall population. The *bars* indicate the number of patients in each group.

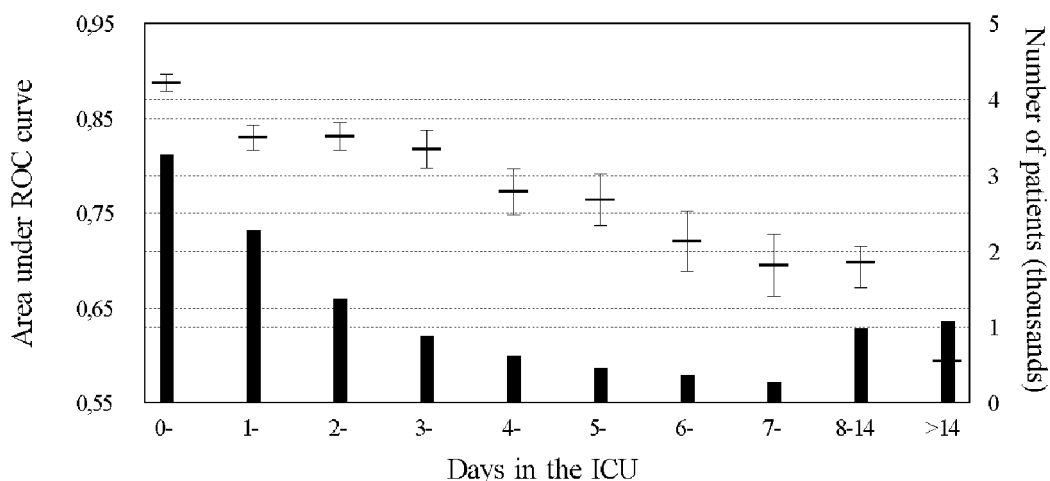


Figure 3. Area under the receiver operating characteristic (ROC) curve and length of stay in the ICU. Results are presented with standard error; *bars* denote the number of patients in each group.

In order to further evaluate the relations between LOS and SAPS II, the database was randomly spliced into development ($n = 7357$ patients) and validation ($n = 3679$ patients) samples, after excluding the patients with a SAPS II less than 7 or greater than 79 ($n = 472$ patients).

After aggregation of SAPS II at 5 points intervals and computing the mean observed LOS in each group, a regression equation was fitted in the development sample and then cross-validated in the validation sample. A model with the square and the cube of SAPS II group as independent variables provided the best fitting, with an R square of 0.97 on the development sample (Figure 4). In the validation sample, the developed model presented a good fitting, with an R square between predicted and observed also of 0.97 (Figure 4).

Figure 4. Predicted versus observed length of stay in the ICU (LOS). Top: development sample (n = 7357 patients); bottom: validation sample (n = 3679 patients). In both panels are displayed the observed LOS (•) and the predicted LOS (*). The *bars* indicate the number of patients in each group.

LOS AND HOSPITAL CHARACTERISTICS

LOS presented significant differences ($p < 0.001$) according to the type of hospital (Table 4), with longer LOS in university hospitals. These differences remained significant after correction for the severity of illness ($p < 0.001$).

The presence of a step-down unit in the same hospital did not reduce LOS. In ICUs located in hospitals with a step-down unit (n = 23), median (interquartile range) LOS was 2.18 days (0.95 to 6.00 days) vs 2.00 days (0.92 to 5.25 days) in ICUs located in hospitals without step-down units (n = 57). This difference was statistically significant ($p < 0.001$). A difference in severity could be demonstrated, with the patients in ICUs located in hospitals without step-down units being more severely ill (SAPS II 35.66 ± 18.66 vs 32.45 ± 17.84 , $p < 0.001$). The differences in LOS remained significant ($p < 0.001$) after correction for the severity of illness. In the ICUs located in hospitals with step-down units, only a median of 5.7 % (interquartile range 0 to 18.7 %, range 0 % to 69.3 %) of the survivors were transferred to step-down units.

Table 4. Length of stay in the ICU (LOS) according to type of hospital.

	Type of hospital			
	< 300 beds	300-500 beds	> 500 beds	University
ICUs	15	18	21	35
Step-down units	2	2	8	11
LOS ^a	2.0 (0.9-4.5)	2.2 (1.0-5.5)	2.0 (0.9-4.9)	2.4 (0.9-6.6)
SAPS-II ^b	31.7 ± 17.9	33.6 ± 17.9	33.54 ± 18.4	34.6 ± 18.4

SAPS II, new Simplified Acute Physiology Score

^a: days, median (interquartile range)

^b: mean \pm standard deviation

DISCUSSION

We analysed the relationships between some relevant patient and hospital characteristics and LOS in the ICU. The results of the study have shown that after correction for the severity of illness LOS was longer in medical patients, in patients admitted from other ICUs/other hospitals or from the ward, in patients with non-operative respiratory or trauma (non-operative or post-operative) diagnoses and in those admitted to university hospitals. These

associations confirm the results of previous studies [6,7,35-37], but are different from others [38].

Rather unexpected was the relation between the presence or absence of step-down units in the hospital and the LOS in the respective ICU, even after correction for the severity of illness. The LOS on ICUs operating in hospitals with step-down units was longer than in those hospitals without step-down units. This goes indeed against the cost-effectiveness argument in favour of the implementation of step-down units, that the presence of step-down units would facilitate the early discharge of patients from the ICU with a consequent reduction on the use of the more expensive beds. [39-41]. However, step-down units can only reduce LOS, allowing early discharge of patients if they work in close contact with the ICU and admit an important part of the survivors. This was not verified in this study, with only a small percentage of the survivors discharged from the ICUs to those units, and this fact can probably explain the observed results. Certainly, more research is needed in how to improve the interrelations between ICUs working at different levels of care.

Overall, non-survivors presented a longer LOS than survivors. When severity is taken into account a complex pattern appears with a linear relation between severity of illness and LOS only in survivors, although variation increases with increasing severity of illness. For non-survivors, as severity of illness increases LOS first increases and then falls. We can submit the hypothesis that this phenomenon can be explained by a combination of early deaths in patients in which markers of physiologic dysfunction did not had time to develop (and consequently with low SAPS II scores) and the rather quick death of patients who were very severely ill. This non-linearity of the relation between severity of illness and LOS in non-survivors can have important consequences in the standardisation of the LOS and in its use as a marker of resources utilisation. Several papers in the literature describe equations to predict LOS, based mainly in the severity of illness [7,35], sometimes with correction for non-linearity [7] and then using the discrepancies between predicted and observed LOS as an indicator of the effectiveness in the use of resources. This approach assumes that the LOS can be used as a proxy for resource utilisation and that severity of illness is a main determinant of LOS. In other words, it is assumed that it is possible to establish a relation between severity of illness and resources use [42]. The non linearity of the relation described is very similar to that found by Oye et al [43] between APACHE II and total Therapeutic Intervention Scoring System (TISS) per admission. This relationship can be expected since non-survivors are responsible for a major part of total resources use in intensive care [44], and challenges the use of linear procedures to relate severity of illness to resources use through the use of standardised LOS, since the pattern of increasing LOS with increasing severity of illness can only be demonstrated in survivors. Our work demonstrated also that this relation can be modelled but we should expect the stability of the equation to be dependent from the proportion of survivors and non-survivors in the target population.

The impact of LOS on the performance of SAPS II deserves attention. In this population, the overall goodness of fit was poor for SAPS II; if the model kept its discriminative power, as measured by the area under ROC curve (0.816 vs 0.86 in ENAS database [4]), calibration was very poor, with significant differences between predicted and observed mortality. When

both parameters were analysed according to LOS, it was seen that SAPS II loses quickly its discriminative power, performing hardly better than chance (area under ROC curve 0.594, standard error 0.018) in patients with a LOS greater than two weeks. Calibration, was also affected, with greater values for the Hosmer-Lemeshow statistic in the groups with longer LOS. This poses serious problems to the customisation of the model, since if for calibration probably easy modifications of the predictive equation will be enough to customise it [17,45,46], when a system loses its discriminative power there is hardly something left to be done [47].

In conclusion, the results of this study have shown that the length of stay of ICU patients is influenced by patient and hospital characteristics other than solely the severity of illness of the patients. The mortality predictive performance of the SAPS-II loses significantly its accuracy when concerning groups of patients with LOS greater than one week. The results of this study could not document that the existence of step-down units do have an beneficial influence on the use of the ICU facilities.

ACKNOWLEDGEMENTS

The authors thank the participating staff of all the ICUs for their full collaboration on data collection efforts.

REFERENCES

- Lemeshow S, Teres D, Klar J, Avrunin JS, Gehlbach SH, Rapoport J. Mortality Probability Models (MPM II) based on an international cohort of intensive care unit patients. *JAMA* 1993;270:2478-86.
- Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med* 1985;13:818-29.
- Knaus WA, Wagner DP, Draper EA, et al. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest* 1991;100:1619-36.
- Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European / North American multicenter study. *JAMA* 1993;270:2957-63.
- Knaus WA, Draper EA, Wagner DP, Zimmerman JE. An evaluation of outcome from intensive care in major medical centers. *Ann Intern Med* 1986;104:410-8.
- Zimmerman JE, Shortell SM, Rousseau DM, et al. Improving intensive care: observations based on organizational case studies in nine intensive care units: a prospective, multicenter study. *Crit Care Med* 1993;21:1443-51.
- Knaus WA, Wagner DP, Zimmerman JE, Draper EA. Variations in mortality and length of stay in Intensive Care Units. *Ann Intern Med* 1993;118:753-61.

- Rowan KM, Kerr JH, Major E, McPherson K, Short A, Vessey MP. Intensive Care Society's APACHE II study in Britain and Ireland - II: Outcome comparisons of intensive care units after adjustment for case mix by the American APACHE II method. *Br Med J* 1993;307:977-81.
- Zimmerman JE, Knaus WA, Judson JA, et al. Patient selection for intensive care: a comparison of New Zealand and United States Hospitals. *Crit Care Med* 1988;16:318-25.
- Sirio CA, Tajimi K, Tase C, et al. An initial comparison of intensive care in Japan and United States. *Crit Care Med* 1992;20:1207-15.
- Wong DT, Crofts SL, Gomez M, McGuire GP, Byrick RJ. Evaluation of predictive ability of APACHE II system and hospital outcome in Canadian intensive care unit patients. *Crit Care Med* 1995;23:1177-83.
- Greenman RL, Schein RMH, Martin MA, et al. A Controlled Clinical Trial of E5 Murine Monoclonal IgM Antibody to Endotoxin in the Treatment of Gram-Negative Sepsis. *JAMA* 1991;266:1097-102.
- Ziegler EJ, Fisher CJ, Sprung CL, et al. Treatment of gram-negative bacteremia and septic shock with ha-1a human monoclonal antibody against endotoxin. A randomized, double-blind, placebo-controlled trial. *N Engl J Med* 1991;324:429-36.
- Knaus WA, Harrell FE, Fisher CJ, et al. The clinical evaluation of new drugs for sepsis. A prospective study design based on survival analysis. *JAMA* 1993;270:1233-41.
- Knaus WA, Harrell FE, LaBrecque JF, et al. Use of predicted risk of mortality to evaluate the efficacy of anticytokine therapy in sepsis. *Crit Care Med* 1996;24:46-56.
- Goldhill DR, Withington PS. The effects of casemix adjustment on mortality as predicted by APACHE II. *Intensive Care Med* 1996;22:415-9.
- Apolone G, D'Amico R, Bertolini G, et al. The performance of SAPS II in a cohort of patients admitted in 99 Italian ICUs: results from the GiViTI. *Intensive Care Med* 1996;22:1368-78.
- Rowan KM, Kerr JH, Major E, McPherson K, Short A, Vessey MP. Intensive Care Society's APACHE II study in Britain and Ireland - I: Variations in case mix of adult admissions to general intensive care units and impact on outcome. *Br Med J* 1993;307:972-7.
- Rowan K. The reliability of case mix measurements in intensive care. *Curr Opin Crit Care* 1996;2:209-13.
- Lemeshow S, Klar J, Teres D, et al. Mortality Probability Models for patients in the intensive care unit for 48 or 72 hours: a prospective, multicenter study. *Crit Care Med* 1994;22:1351-8.
- Zimmerman JE, Wagner DP, Draper EA, Knaus WA. Improving intensive care unit discharge decisions: supplementary physician judgement with predictions of next day risk for life support. *Crit Care Med* 1994;22:1373-84.
- Chang RW, Jacobs S, Lee B. Use of APACHE II severity of disease classification to identify intensive-care-unit patients who would not benefit from total parenteral nutrition. *Lancet* 1986;1483-6.
- Chang RWS, Jacobs S, Lee B, Pace N. Predicting outcome among intensive care unit patients. *Crit Care Med* 1988;16:34-42.
- Chang RWS, Jacobs S, Lee B. Predicting outcome among intensive care unit patients using computerised trend analysis of daily APACHE II scores corrected for organ system failure. *Intensive Care Med* 1988;14:558-66.

- Rogers J, Fuller HD. Use of daily Acute Physiology and Chronic Health Evaluation (APACHE) II scores to predict individual patient survival rate. *Crit Care Med* 1994;22:1402-5.
- Vincent J-L, Bihari DJ, Suter PM, et al. The prevalence of nosocomial infection in intensive care units in Europe. Results of the European Prevalence of Infection in Intensive Care (EPIC) Study. *JAMA* 1995;274:639-44.
- Maher ER, Robinson KN, Scoble JE, et al. Prognosis of critically ill patients with acute renal failure: APACHE II score and other predictive factors. *Q J Med* 1989;72:857-66.
- Cerra FB, Negro F, Abrams J. APACHE II score does not predict multiple organ failure or mortality in post-operative surgical patients. *Arch Surg* 1990;125:519-22.
- Hopfel AW, Taaffe CL, Herrmann VM. Failure of APACHE II alone as a predictor of mortality in patients receiving total parenteral nutrition. *Crit Care Med* 1989;17:414-7.
- Niskanen M, Kari A, Nikki P, et al. Acute Physiology and Chronic Health Evaluation (APACHE II) and Glasgow coma scores as predictors of outcome from intensive care after cardiac arrest. *Crit Care Med* 1991;19:1465-73.
- Management of Intensive Care. Guidelines for better use of resources. Reis Miranda D, Williams A, Loirat P, eds. Dordrecht/Boston/London: Kluwer Academic Publishers, 1990.
- Hanley J, McNeil B. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29-36.
- Lemeshow S, Hosmer DW. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol* 1982;115:92-106.
- Hosmer DW, Lemeshow S. *Applied logistic regression*. New York, NY: John Wiley & Sons, Inc., 1989.
- Rapoport J, Teres D, Lemeshow S, Gehlbach S. A method for assessing the clinical performance and cost-effectiveness of intensive care units: a multicenter inception cohort study. *Crit Care Med* 1994;22:1385-91.
- Zimmerman JE, Shortell SM, Knaus WA, et al. Value and cost of teaching hospitals: a prospective, multicenter, inception cohort study. *Crit Care Med* 1993;21:1432-42.
- Shortell SM, Zimmerman JE, Rousseau DM, et al. The performance of intensive care units: does good management make a difference ? *Med Care* 1994;32:508-25.
- Dragsted L, Qvist J. Outcome from intensive care. I. A 5-year study of 1308 patients: methodology and patient population. *Eur J Anaesthesiol* 1989;6:23-37.
- Ryan DW. Rationing intensive care. High dependency units may be the answer [letter]. *Br Med J* 1995;310:1010-1.
1. Bion J. Rationing intensive care. *Br Med J* 1995;310:682-3.
 2. Peacock JE, Edbrooke D. Data from high dependency unit supports their effectiveness [letter]. *Br Med J* 1995;310:1413.
 3. Rapoport J, Teres D, Lemeshow S, Avrunin JS, Haber R. Explaining variability of cost using a severity of illness measure for ICU patients. *Med Care* 1990;28:338-48.
 4. Oye RK, Bellamy PF. Patterns of resource consumption in medical intensive care. *Chest* 1991;99:695-89.

5. Atkinson S, Bihari D, Smithies M, Daly K, Mason R, McColl I. Identification of futility in intensive care. *Lancet* 1994;344:1203-6.
6. Le Gall J-R, Lemeshow S, Leleu G, et al. Customized probability models for early severe sepsis in adult intensive care patients. *JAMA* 1995;273:644-50.
7. Zhu B-P, Lemeshow S, Hosmer DW, Klarm J, Avrunin J, Teres D. Factors affecting the performance of the models in the Mortality Probability Model and strategies of customization: a simulation study. *Crit Care Med* 1996;24:57-63.
8. Miller ME, Hui SL. Validation techniques for logistic regression models. *Stat Med* 1991;10:1213-26.